A GENERALIZATION OF THE FRANK-WOLFE THEOREM

BY

ANDRE F. PEROLD

TECHNICAL REPORT 77-18

July 1977

Department of
Operations
Research

Stanford
University

Stanford
California
94305

A GENERALIZATION OF THE FRANK-WOLFE THEOREM

by

Andre F. Perold

TECHNICAL REPORT 77-18

July 1977

DEPARTMENT OF OPERATIONS RESEARCH

Stanford University
Stanford, California

Abstract

    The Frank-Wolfe Theorem states that a quadratic function, bounded below on a nonempty polyhedral convex set, attains its infimum there. This paper gives sufficient conditions under which a function either attains its infimum on a nonempty polyhedral convex set or is unbounded below on some halfline of that set. Quadratic functions are shown to satisfy these sufficient conditions.

## 1. Introduction

The existence theorem for quadratic programming states that a quadratic function $Q$ bounded below on a nonempty polyhedral convex set $C$ attains its infimum there. This result was first proved in 1956 by Frank and Wolfe [1]. Alternative proofs have since been given by Collatz and Wetterling [2] (for the case when $Q$ is convex), Eaves [3] and Blum and Oettli [4].

Eaves [3] also improved on this result by showing that if $Q$ does not attain its infimum on $C$, then $Q$ must be unbounded below on some halfline contained in $C$. This was first claimed, but not proved, by Dennis in 1959 [5].

In [1], the example $Q(x_1, x_2) = x_1^2 + (1-x_1 x_2)^2$ is given to show that these results do not hold in general for higher order polynomials. ($Q$ here does not attain its infimum, zero, in the plane.) This leads one to ask: what is so special about a quadratic; and also perhaps, what is so special about polyhedral convex sets. Unfortunately, the proofs in [1], [2], [3] and [4] are specifically tailored to the quadratic and polyhedral case, and shed little light on the answers to these questions.

In this paper we shall deal with the first question. Let $\mathcal{F}$ be the class of all continuous functions $f: \mathbb{R}^n \to \mathbb{R}^1$ (some n) such that for any polyhedral convex set $C \subseteq \mathbb{R}^n$, f either attains its infimum or $C$ or is unbounded below on some halftime contained in $C$. We shall define a class of functions $\mathcal{A}$ which, in particular, contains all quadratic functions, and shall show that $\mathcal{A}$ is contained (strictly) in $\mathcal{F}$.

## 2. Notation and Definitons

Let $\underline{n}$ denote $\{1, 2, \ldots, n\}$. For $x \in R^n$, $\alpha \subseteq \underline{n}$, let $x_\alpha \in R^k$ denote

$$(x_{\alpha_1}, \ldots, x_{\alpha_k})^T \quad \text{where} \quad \alpha = \{\alpha_1, \ldots, \alpha_k\}, \ \alpha_1 < \ldots < \alpha_K$$

1

For $A \in R^{m \times n}$, $\alpha \subseteq \underline{m}$, $\beta \subseteq \underline{n}$, let $A_{\alpha \bullet}$ denote the row submatrix of $A$ whose rows are indexed by $\alpha$; let $A_{\bullet \beta}$ denote the column submatrix of $A$ whose columns are indexed by $\beta$; let $A_{\alpha \beta}$ denote $(A_{\alpha \bullet})_{\bullet \beta}$ .

The symbol $//$ will be used to denote the end of a proof.

A function $f : R^n \to R$ is said to be <u>norm-coercive</u> if $\lim\limits_{k \to \infty} |f(x_k)| = \infty$ for every sequence $\{x_k\} \subseteq R^n$ such that $\lim\limits_{k \to \infty} \| x_k \| = \infty$.

## 3. The class $\mathscr{H}$ and some examples

We define $\mathscr{H}$ by induction and the number of variables.

<u>3.1 Definition</u>: Let $f : R^n \to R^1$ be continuous. Then $f \in \mathscr{H}$ if there exists a nonsingular matrix $Q$, and a partition of $\underline{n}$, $\underline{n} = \alpha \cup \beta$, such that $g : R^n \to R$ defined by $g(x) = f(Qx)$ satisfies

(i) if $\beta \neq \phi$ then $g(x_\alpha, \cdot)$ is a concave function

(ii) if $\beta = \phi$ then $g$ is norm-coercive

(iii) if $\alpha \neq \phi$ and $\beta \neq \phi$ then for all $A$ and $b$ of appropriate dimensions, the function $h(\cdot)$, defined by $h(x_\alpha) = g(x_\alpha, Ax_\alpha + b)$, is a member of $\mathscr{H}$.

Note that for $n = 1$, condition (iii) is trivally satisfied since either $\alpha$ or $\beta$ must be empty. Hence a function of one variable, $f$, is in $\mathscr{H}$ if $f$ is concave or if $f$ is norm-coercive and continuous.

Note also that the induction step in the defn of $\mathscr{H}$ is in condition (iii).

### 3.2 Examples

(i) All continuous norm-coercive functions are in $\mathscr{H}$.

Set $Q = I$, $\alpha = \underline{n}$, $\beta = \phi$ in Definition 3.1

(ii) All concave functions are in $\mathscr{H}$.

Set $Q = I$, $\alpha = \phi$, $\beta = \underline{n}$ in Definition 3.1

(iii) All quadratic functions are in $\mathscr{H}$.

2

We shall prove this by induction on $n$, the number of variables. Let

$$f(x) = c^T x + \frac{1}{2} x^T D x, \quad x \in R^n.$$

If $n = 1$ then $f$ is either concave or $f(x) \to \infty$ as $|x| \to \infty$, and we are done.

Suppose the result is true for quadratic functions of $n-1$ variables or fewer. Without loss of generality we may assume that $D$ is symmetric. Hence there is an orthogonal matrix $Q$ such that

$$Q^T D Q = \text{Diag} (\lambda_1, \ldots, \lambda_n)$$

where the $\lambda_i$ are the real eigenvalues of $D$.

Set $d = Q^T c$, and $g(x) = f(Qx)$. Then

$$g(x) = \sum_{\lambda_i > 0} (\lambda_i x_i^2 + d_i x_i) + \sum_{\lambda_i \leq 0} (\lambda_i x_i^2 + d_i x_i)$$

Let $\alpha = \{i : \lambda_i > 0\}$, $\beta = \{i : \lambda_i \leq 0\}$. Then $\alpha$ and $\beta$ partition $\underline{n}$, and $g(x)$ clearly satisfies conditions (i) and (ii) of Definition 3.1.

Also, since for any $A$ and $b$, $g(x_\alpha, Ax_\alpha + b)$ is a quadratic function in fewer variables, the induction hypothesis applies and condition (iii) of Definition 3.1 is established.

(iv) Any function of the form

$$f(u,v) = [u^T B u]^k - (v^T D v)^m$$

where $B$ is positive definite, $D$ is positive semidefinite and $k > m \geq 0$, is in $\mathcal{H}$.

Set $Q = I$, and identify $x_\alpha$ with $u$ and $x_\beta$ with $v$. Since $B$ is positive definite

$$x_\alpha^T B x_\alpha \geq \xi ||x_\alpha||^2 \quad \forall \, x_\alpha \tag{1}$$

where $\xi > 0$ is the smallest eigenvalue of $B$.

3

Since D is positive semi-definite, and the function $t^m$ is nondecreasing in $t \geq 0$, $(x_\beta^T D x_\beta)^m$ is convex in $x_\beta$.

Hence conditions (i) and (ii) of Definition 3.1 are satisfied.

Further, since D is positive semi-definite,

$$0 \leq x_\beta^T D x_\beta \leq \eta \| x_\beta \|^2 \quad \forall x_\beta$$

where $\eta \geq 0$ is the largest eigenvalue of D. Thus for any A and b

$$0 \leq (Ax_\alpha + b)^T D(Ax_\alpha + b)$$

$$\leq \eta \| Ax_\alpha + b \|^2$$

$$\leq \eta (\mu \| x_\alpha \| + \| b \|)^2 \qquad (2)$$

where $\mu$ is the largest row norm of A.

Combining (1) and (2) we obtain

$$f(x_\alpha, Ax_\alpha + b) \geq \xi^k \| x_\alpha \|^{2k} - \eta^m (\mu \| x_\alpha \| + \| b \|)^{2m}$$

The right hand side of this inequality is a polynomial in $\| x_\alpha \|$ with leading coefficient $\xi^k > 0$. Hence $f(x_\alpha, Ax_\alpha + b) \to \infty$ as $\| x_\alpha \| \to \infty$. This establishes condition (iii) of Definition 3.1 by an application of example (i).

Remark: Example (iv) shows that the class $\mathcal{G}$ is indeed interesting, that is, $\mathcal{G}$ consists of functions to which the Frank-Wolfe theorem as it stands is not applicable.

4

## 4. Preliminary results.

Let  $C$  be a nonempty polyhedral convex set of the form

$$C = \{x \in R^n : Ax \geq b\}$$

where  $A \in R^{mxn}$ .

**4.1 Definition:** Let  $x \in C$ , and let  $\gamma = \{i : A_{i.}\, x = b_i\}$ . Then  $x$  is called a pseudo-extreme point of  $C$  if  $x$  uses linearly independent columns of  $A_{\gamma.}$ , i.e. if  $\delta = \{j: x_j \neq 0\}$  then  $A_{\gamma\delta}$  has full column rank.

For convenience, when  $A$  is the zero matrix, define the origin to be the pseudo-extreme point of  $C$ .

One can easily show that  $C$  always has pseudo-extreme points. Furthermore, it follows by Lemma 4.3 (stated below) that the extreme points of  $C$ , when they exist, are pseudo-extreme points of  $C$ .

Geometrically, the pseudo-extreme points of  $C$  are the extreme points of all the sections of  $C$  at  $x_\alpha = 0$  where  $\alpha$  ranges over all subsets (including the empty set) of  $\underline{n}$

In the following examples, the pseudo-extreme points are marked with asteriscs. The set  $C$  is the shaded area.

(i)   $C = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right\}$
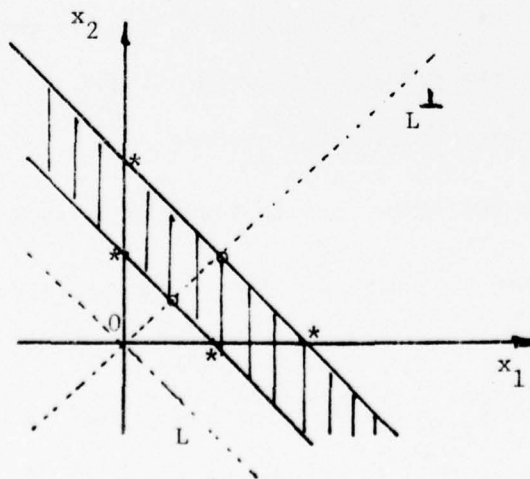


Fig. 1

(ii) $\quad C = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$
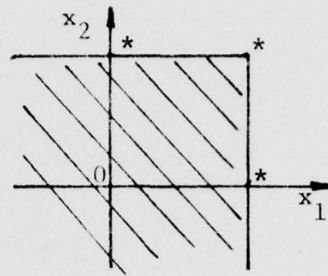


Fig. 2

The following theorem shows that $C$ has a representation in terms of its pseudo-extreme points.

4.2 Theorem:  $C$ has a representation

$$C = \{s + \mu t : s \in S, \ t \in T, \ \mu \geq 0\}$$

where $S$ is the convex hull of the pseudo-extreme points of $C$, and $T$ is the intersection of a polyhedral convex cone with the unit sphere.

To prove this result, we require 3 lemmas.

4.3 Lemma:  Let $x \in C$ and let $\gamma = \{i : A_i \cdot x = b_i\}$. Then $x$ is an extreme point of $C$ iff the rank of $A_{\gamma \cdot}$ is $n$.

4.4 Lemma:  If $C$ has extreme points then $C$ has the representation

$$C = \{s + \mu t : s \in S, \ t \in T, \ \mu \geq 0\} \quad \text{where}$$

$S$ is the convex hull of the extreme points of $C$ and $T$ is the intersection of a polyhedral convex cone with the unit sphere.

The proofs of these two lemmas can be found in Goldman [6].

In the following, let $x^+$ and $x^-$ denote the positive and negative parts of $x$, respectively.

4.5 Lemma:  Let $\overline{C} = \left\{ \begin{pmatrix} u \\ v \end{pmatrix} \geq 0 : Au - Av \leq b \right\}$. Then $x$ is a pseudo-extreme point of $C$ if and only if $\begin{pmatrix} x^+ \\ x^- \end{pmatrix}$ is an extreme point of $\overline{C}$.

6

<u>Proof</u>: Let $x$ be a pseudo-extreme point of $C$. Let $\gamma = \{i : A_{i\cdot} x = b_i\}$.

Then since $x$ uses linearly independent columns of $A_{\gamma\cdot}$, $\begin{pmatrix} x^+ \\ x^- \end{pmatrix}$ uses linearly independent columns, say $\delta \subseteq \underline{2n}$, of the partitioned matrix $(A_{\gamma\cdot}, -A_{\gamma\cdot})$. Let $B = (A, -A)$ and let $I$ be the identity matrix of dimension $2n$. Then $\overline{C}$ can be written as

$$\overline{C} = \{ \begin{pmatrix} u \\ v \end{pmatrix} : \overline{A} \begin{pmatrix} u \\ v \end{pmatrix} \leq \begin{pmatrix} b \\ 0 \end{pmatrix} \}$$

where

$$\overline{A} = \begin{pmatrix} B \\ I \end{pmatrix}$$

Thus the row submatrix of $\overline{A}$ corresponding to the constraints which $\begin{pmatrix} x^+ \\ x^- \end{pmatrix}$ satisfies with equality is

$$\begin{pmatrix} B_{\gamma\cdot} \\ \\ I_{\overline{\delta}\cdot} \end{pmatrix}$$

where $\overline{\delta}$ denotes $\underline{2n} \sim \delta$

Rewrite this matrix as

$$\begin{pmatrix} B_{\gamma\delta} & B_{\gamma\overline{\delta}} \\ \\ 0 & I_{\overline{\delta}\overline{\delta}} \end{pmatrix}$$

Since $B_{\gamma\delta}$ and $I_{\overline{\delta}\overline{\delta}}$ both have full column rank it follows that

$$\text{rank} \begin{pmatrix} B_{\gamma\cdot} \\ \\ I_{\overline{\delta}\cdot} \end{pmatrix} = 2n$$

By lemma 4.3 $\begin{pmatrix} x^+ \\ x^- \end{pmatrix}$ is an extreme point of $\overline{C}$, and the first implication is established.

7

The converse follows easily along similar lines.　　　　//

Proof of Theorem 4.2:　Let $\overline{C}$ be as in the statement of Lemma 4.5.
Since  C  has pseudo-extreme points, it follows by Lemma 4.5 that $\overline{C}$ has
extreme points.

Thus by Lemma 4.4 $\overline{C}$ has a representation

$$\overline{C} = \{ \begin{pmatrix} s' \\ s'' \end{pmatrix} + \mu \begin{pmatrix} t' \\ t'' \end{pmatrix} : \begin{pmatrix} s' \\ s'' \end{pmatrix} \in \overline{S}, \begin{pmatrix} t' \\ t'' \end{pmatrix} \in \overline{T}, \mu \geq 0 \}$$

where $\overline{S}$ is the convex hull of the extreme points of $\overline{C}$, and $\overline{T}$ is the
intersection of a polyhedral convex cone with the unit sphere.

Now　　　C  can be written as

$$C = \{ u-v : \begin{pmatrix} u \\ v \end{pmatrix} \in \overline{C} \}$$

Hence

$$C = \{ (s' - s'') + \mu(t'-t'') : \begin{pmatrix} s' \\ s'' \end{pmatrix} \in \overline{S}, \begin{pmatrix} t' \\ t'' \end{pmatrix} \in \overline{T}, \mu \geq 0 \}$$

Set

$$S = \{ s'-s'' : \begin{pmatrix} s' \\ s'' \end{pmatrix} \in \overline{S} \}$$

Then by Lemma 4.5  S  is the convex hull of the pseudo-extreme points of  C.

Set

$$T = \{ \frac{t' - t''}{\| t' - t'' \|} : \begin{pmatrix} t' \\ t'' \end{pmatrix} \in \overline{T}, t' - t'' \neq 0 \}$$

Then

$$C = \{ s + \mu t : s \in S, t \in T, \mu \geq 0 \}$$

as required.

Theorem 4.6:　Let  $g : C \rightarrow \mathbb{R}$  be concave.  Then either  g  attains its infimum
at a pseudo-extreme point of  C  or  g  is unbounded below on some halfline of  C.


In 1961 Hirsch and Hoffman [7] proved a similar theorem using a different
representation of  C.  They decomposed  C  as

$$C = M \oplus L$$

8

where  L  is a linear subspace and  M  is the  $L^{\perp}$ - section of  C, and showed

that a concave function bounded below on  C  attains its infimum at an extreme

point of  M.  In fig. 1 the extreme points of  M  are circled.  Note that they

are not the same as the pseudo-extreme points of  C.

Proof of Theorem 4.6:

Let  $\{p_1, \ldots, p_k\}$  be the set of pseudo-extreme points of  C.

Let

$$g(p_m) = \min_{1 \leq i \leq k} g(p_i)$$

Let  S  and  T  be as in Theorem 4.2.

For any  $s \in S$  $\exists$  $\lambda_i \geq 0$, $i = 1, \ldots, k$,  $\sum_{i=1}^{k} \lambda_i = 1$  such that

$$s = \sum_{i=1}^{k} \lambda_i p_i$$

By the concavity of  g,

$$g(s) = g\left(\sum_{i=1}^{k} \lambda_i p_i\right) \geq \sum_{i=1}^{k} \lambda_i g(p_i)$$

Hence  $g(s) \geq g(p_m)$  $\forall s \in S$. (1)

Now suppose that  g  does not attain its infimum at a pseudo-extreme point of

C.

Then     by (1) $\exists$ $x \in C \setminus S$  with

$$g(x) < g(p_m)$$

By Theorem 4.2  x  can be written as

$$x = s + \mu t$$

for some  $s \in S, t \in T, \mu \geq 0$.

Since  $x \notin S, \mu > 0$.

Further, the half-line

$$H = \{s + \xi t : \xi \geq \mu\}$$

is contained in C.

Now $\quad \forall \; \xi \geq \mu$

$$s + \mu t = (1 - \frac{\mu}{\xi})\, s + (\frac{\mu}{\xi})(s + \xi t) \, .$$

Since g is concave

$$g(s + \mu t) \geq (1 - \frac{\mu}{\xi})g(s) + (\frac{\mu}{\xi})g(s + \xi t) \, .$$

Upon rearrangement we get

$$g(s + \xi t) \leq g(s) - (\frac{g(s) - g(s + \mu t)}{\mu})\xi \, .$$

Since

$$g(s) - g(s + \mu t) \geq g(pm) - g(x) > 0$$

it follows that

$$g(s + \xi t) \to -\infty \quad \text{as} \quad \xi \to \infty.$$

i.e. g is unbounded below on H.

$\hfill //$

The following result is of interest in its own right.

<u>Theorem 4.7</u>: Let $g : R^n \to R$, $n \geq 2$, be continuous and norm-coercive. Then g is either bounded above or bounded below.

The proof of this result uses the concept of path connectedness. A set $D \subset R^n$ is said to be <u>path connected</u> if for any $x, y \in D$ there is a continuous mapping $p : [0,1] \to D$ such that $p(0) = x$ and $p(1) = y$.

Let $K_r$ denote the closed ball of radius r, and $\overline{K}_r$ its complement. For $n \geq 2$ $\overline{K}_r$ is path connected.

Proof of Theorem 4.7:  By assumption on  g, there exists  r > 0  such that

$$x \in \bar{K}_r \Rightarrow |g(x)| \geq 1. \qquad (1)$$

We shall show that either  $g(x) \geq 1$  $\forall$  $x \in \bar{K}_r$  or  $g(x) \leq -1$  $\forall$  $x \in \bar{K}_r$.
If not,  $\exists$  x, y $\in \bar{K}_r$  $\ni$  $g(x) \geq 1$  and  $g(y) \leq -1$.  Since  $K_r$  is path
connected there is a continuous mapping  $p : [0,1] \to \bar{K}_r \ni p(0) = x \quad p(y) = 1$

Since  g  is continuous, the mapping  $h : [0,1] \to R$  defined by

$$h(t) = g(p(t))$$

is continuous.  Moreover  $h(0) \geq 1$  and  $h(1) \leq -1$.  By the intermediate value
theorem there exists  $t_0 \in [0,1]$  such that  $h(t_0) = 0$,

   i.e.   $g(p(t_0) = 0$

Since  $p(t_0) \in \bar{K}_r$  this contradicts (1).
Hence  g  is either bounded above or below on  $\bar{K}_r$.  Since  $K_r$  is compact and
g  is continuous, g  is bounded on  $K_r$.                                           //

   The following theorem is obtained as a consequence of the above result.
Theorem 4.8:  Let  $g : R^n \to R$  be continuous and norm-coercive.
Let  P  be any nonempty closed convex set.  Then  g  either attains its infimum
on  P  or is unbounded below on some half-line contained in  P.
Proof:  If  g  is bounded below on  P, then

$$\|x\| \to \infty, \ x \in P \Rightarrow g(x) \to \infty.$$

Thus for  k  sufficiently large the set

$$R = \{x : g(x) \leq k\} \cap P$$

is nonempty and compact.

Since  g  is continuous  g  attains its infimum on  R.  The infimum of  g  on  R
is equal to the infimum of  g  on  P  and we are done.

11

Suppose  g  is unbounded below on  P.  Then since  g  is continuous, there is
a sequence  $\{x_n\} \subseteq P \ni g(x_n) \to -\infty$  and  $\|x_n\| \to \infty$.

If  n = 1, one of the half-lines  $P \cap [-\infty, 0]$,  $P \cap [0, \infty)$  will do.

If  $n \geq 2$, it follows by Theorem 4.7 that

$$\| x \| \to \infty, \ x \in P \Rightarrow g(x) \to -\infty.$$

Since any unbounded convex set contains half-lines, any half-line in  P  will do.   //

4.9 Lemma:  Let  $P(b) = \{x : A x \leq b\}$.  Then there is a finite collection
$\{(\gamma_i, \delta_i) \ i = 1, \ldots, \ell\}$  where  $\gamma_i \subseteq \underline{m}$,  $\delta_i \subseteq \underline{n}$  such that
$\forall b$, x is a pseudo-extreme point of  P(b) $\Rightarrow$  there exists  $i \in \underline{\ell}$  such that

    (i)  $A_{\gamma_i \delta_i} x_{\delta_i} = b_{\gamma_i}$

    (ii)  $A^{-1}_{\gamma_i \delta_i}$  exists

    (iii)  $x_{\bar{\delta}_i} = 0$     where  $\bar{\delta}_i = \underline{n} \sim \delta_i$

The proof is immediate from Definition 4.1 and will be omitted.

## 5. Proof of the main result: $\mathcal{H} \subset \mathcal{F}$

The proof is by induction on the number of variables.

Let  C  be any non-empty polyhedral convex set defined by

$$C = \{x \in R^n : A x \leq b\}$$

where  $A \in R^{m \times n}$.

For  n = 1, if  $f \in \mathcal{H}$ then  f  is either concave or  $|f(x)| \to \infty$  as  $|x| \to \infty$.
In the former case, Theorem 4.6 applies.  In the latter case, Theorem 4.8 applies.

Suppose the result holds for all  k < n.  Let  $f \in \mathcal{H}$, $f : R^n \to R$.
Let  Q, $\alpha$ and  $\beta$  be as in definition 3.1.  Since  Q  is nonsingular, it
suffices to prove the result for  g  defined by

12

$$g(x) = f(Qx)$$

on the set

$$\overline{C} = \{x : \overline{A} \, x \le b\}$$

where $\overline{A} = AQ$ .

If either $\alpha = \phi$ or $\beta = \phi$, then $g$ is either concave or norm-coercive and we can use Theorem 4.6 or 4.8.

So assume $\alpha \neq \phi$, $\beta \neq \phi$. Let $\{x^k\} = \{(x_\alpha^k, x_\beta^k)\}$ be a minimizing sequence for $g$ in the sense that $g(x^k) \searrow - \infty$ if $g$ is unbounded below on $\overline{C}$, or else

$$g(x^k) \searrow \inf \{g(x) : x \in \overline{C}\} = \rho > - \infty \quad \text{(say)}.$$

Now for each $k$, hold $x_\alpha^k$ fixed and solve

$$\min \quad g(x_\alpha^k, u_\beta)$$

subject to $u_\beta \in \overline{C}_k \triangleq \{y : \overline{A}_{\cdot\alpha} x_\alpha^k + \overline{A}_{\cdot\beta} y \le b\}$. Since $g(x_\alpha^k, \cdot)$ is concave, we can apply theorem 4.6:

If $g(x_\alpha^k, \cdot)$ is unbounded below on $\overline{C}_k$ then there is a half line $\{s + \mu t : \mu \ge 0\} \subseteq \overline{C}_k$ such that

$$g(x_\alpha^k, s + \mu t) \to - \infty \quad \text{as} \quad \mu \to \infty.$$

i.e. $g$ is unbounded below on the half-line

$$\left\{ \begin{pmatrix} x_\alpha^k \\ s \end{pmatrix} + \mu \begin{pmatrix} 0 \\ t \end{pmatrix} : \mu \ge 0 \right\} \subseteq \overline{C}$$

and the result is proved.

Otherwise, $\forall k$, $g(x_\alpha^k, \cdot)$ attains its infimum at a pseudo-extreme point $u_\beta^k \in \overline{C}_k$. By Lemma 4.9 there exists $(\gamma_1, \delta_1)$ such that

13

$$\overline{A}_{\gamma_i \delta_i} \, u^k_{\delta_i} = b_{\gamma_i} - \overline{A}_{\gamma_i \alpha} \, x^k_{\alpha}$$

$$\overline{A}^{-1}_{\gamma_i \delta_i} \quad \text{exists} \tag{1}$$

$$u^k_{\overline{\delta}_i} = 0 \quad \text{where} \quad \overline{\delta}_i = \beta \sim \delta_i$$

Since the pair $(x^k_{\alpha}, u^k_{\beta})$ is feasible

$$\overline{A}_{\overline{\gamma}_i \beta} \, u^k_{\beta} \leq b_{\overline{\gamma}_i} - \overline{A}_{\overline{\gamma}_i \alpha} \, x^k_{\alpha} \tag{2}$$

where $\overline{\gamma}_i = \underline{m} \sim \gamma_i$

Since there are only finitely many such $(\gamma_i, \delta_i)$ there is an infinite subsequence $K$ such that some $(\gamma_i, \delta_i) = (\gamma, \delta)$ (say) is repeated for all $k \in K$. Using (1) to eliminate $u^k_{\beta}$ from (2) we get, $\forall k \in K$

$$u^k_{\beta} = d - D x^k_{\alpha}$$

$$B x^k_{\alpha} \leq c$$

where

$$d = \begin{pmatrix} d_{\delta} \\ d_{\overline{\delta}} \end{pmatrix} = \begin{pmatrix} \overline{A}^{-1}_{\gamma\delta} \, b_{\gamma} \\ 0 \end{pmatrix}$$

$$D = \begin{pmatrix} D_{\delta} \cdot \\ D_{\overline{\delta}} \cdot \end{pmatrix} = \begin{pmatrix} \overline{A}^{-1}_{\gamma\delta} \, \overline{A}_{\gamma\alpha} \\ 0 \end{pmatrix}$$

$$c = b_{\overline{\gamma}} - \overline{A}_{\overline{\gamma}\beta} \, d$$

$$B = \overline{A}_{\overline{\gamma}\alpha} - \overline{A}_{\overline{\gamma}\beta} \, D$$

14

Now since $\{(x_\alpha^k, x_\beta^k)\}$ is a minimizing sequence for $g(\cdot)$ on $\overline{C}$, it follows that $\{(x_\alpha^k, u_\beta^k)\}$ is a minimizing sequence for $g(\cdot)$ on $\overline{C}$ since by definition of $u_\beta^k$,

$$g(x_\alpha^k, u_\beta^k) \leq g(x_\alpha^k, x_\beta^k)$$

But $(x_\alpha^k, u_\beta^k) = (x_\alpha^k, d - Dx_\alpha^k)$ $\forall k \in K$. Hence $\{x_\alpha^k\}_{k \in K}$ is a minimizing sequence for $h(\cdot)$ on $\overline{\overline{C}}$ where

$$h(z) \overset{\Delta}{=} g(z, d - Dz)$$

$$\overline{\overline{C}} = \{z : Bz \leq c\}$$

We can now apply the induction hypothesis to $h(\cdot)$ on $\overline{\overline{C}}$ using condition (iii) of Definition 3.1 as follows:

If $g$ is bounded below on $\overline{C}$ then

$$g(x_\alpha^k, x_\beta^k) \searrow \rho \qquad \text{as} \quad k \to \infty$$

so that

$$g(x_\alpha^k, d - Dx_\alpha^k) = h(x_\alpha^k) \searrow \rho \qquad \text{as} \quad k \to \infty, k \in K.$$

Hence there exists $\overline{x}_\alpha \in \overline{\overline{C}}$ such that $h(\overline{x}_\alpha) = \rho$ i.e. $g(\overline{x}_\alpha, d - \overline{D}x_\alpha) = \rho$, and we are done.

Otherwise $g(x_\alpha^k, x_\beta^k) \searrow -\infty$ so that $h(z_\alpha^k) \searrow -\infty$, $k \in K$, i.e. $h(\cdot)$ is unbounded below on $\overline{\overline{C}}$. Hence there exists a half-line $\{s + \mu t : \mu \geq 0\} \subseteq \overline{\overline{C}}$ such that

$$h(s + \mu t) \searrow -\infty \quad \text{as} \quad \mu \to \infty$$

i.e. $g(\cdot)$ is unbounded below on the half-line

$$\left\{ \begin{pmatrix} s \\ d - Ds \end{pmatrix} + \mu \begin{pmatrix} t \\ -Dt \end{pmatrix} : \mu \geq 0 \right\}$$

in $\overline{C}$. This completes the proof. $\qquad\qquad //$

15

## 6. Concluding Remarks

We might ask questions about the structures of $\mathscr{S}$ and $\mathscr{F}$.

(i) Neither is closed under addition

e.g. Let $f(x) = e^{x^2} + e^{-x}$, $g(x) = -e^{x^2}$. Then both $f$ and $g \in \mathscr{H}$ but $f + g \notin \mathscr{F}$. This example also shows that $\mathscr{S}$ is properly contained in $\mathscr{F}$.

(ii) $\mathscr{F}$ is closed under arbitrary linear linear transformations of the variables whereas $\mathscr{S}$ is not.

To prove the first assertation, let $f \in \mathscr{F}$, and let $C = \{x : Ax \leq b\}$. We must show that $f(Bx + d)$ has the required properties on $C$. Let $E = \{(x,y) : Ax \leq b, y = Bx + d\}$. Since $E$ is a polyhedral convex set, its projection onto $y$ space, $\pi E$ is a polyhedral convex set.

If

$$\inf\{f(Bx + d) : x \in E\} = \rho > -\infty$$

then

$$\inf\{f(y) : y \in \pi E\} = \rho.$$

Since $f \in \mathscr{F}$, there exists $\overline{y} \in \pi E$ such that $f(\overline{y}) = \rho$
By definition of $\pi E$ there exists $\overline{x} \in C$ such that $\overline{y} = B\overline{x} + d$. Hence $f(B\overline{y} + d) = \rho$, and we are done.

If $f(Bx + d)$ is unbounded below then $f(y)$ is unbounded on $\pi E$.
Hence $\exists$ half-line $H = \{s + \mu t : \mu \geq 0\} \subseteq \pi E$ such that

$$f(y) \to -\infty \text{ as } ||y|| \to \infty, \; y \in H.$$

Define the linear map $h : C \to R^n$ by

$$h(x) = Bx + d.$$

Then by linearity of $h(\cdot)$, $h^{-1}(H)$ is an unbounded convex subset of $C$, and hence contains a half-line, $G$.
Thus $f(Bx + d) \to -\infty$ as $||x|| \to \infty$, $x \in G$, as required.

16

To show $\mathcal{H}$ is not closed under arbitrary linear transformations of the variables, consider

$$f(x) = e^{x^2} \in \mathcal{H}.$$

The function $f(y + z) = e^{(y+z)^2}$ is concave neither in $y$ nor $z$, and along the ray $y = -z$, $\left|f(y+z)\right| \not\to \infty$, i.e. $f(y+z) \notin \mathcal{H}$.

Note that $\mathcal{H}$ is however closed under nonsingular linear transformations of the variables. This follows by definition of $\mathcal{H}$.

## Acknowledgement

## Bibliography

1. Frank, M. and Wolfe, P. "An algorithm for Quadratic Programming", Naval
   Research Logistics Quarterly, Vol. 3 (1956) p. 95-110.

2. Collatz, L. and Wetterling, W. "Optimization Problems", Springer (1975).

3. Blum, E. and Oettli, W. "Direct Proof of the Existence Theorem for Quadratic
   Programming". Operations Research (20) (1972) p. 165-167.

4. Eaves, B.C. "On Quadratic Programming", Management Science, Vol. 17 #11
   July 1971 p. 698-711.

5. Dennis, J.B. "Mathematical Programming and Electrical Networks" Wiley
   and Tech Press of MIT. (1959).

6. Goldman, A.J. "Resolution and Separation Theorems for Polyhedral Convex
   Sets" in "Linear Inequalities and Related Systems", Kuhn & Tucker
   (Eds) (1956) p. 41-51.

7. Hirsch, W.M. and Hoffman, A.J. "Extreme Varieties, Concave Functions and
   the Fixed Charge Problem". Comm on Pure and Appl. Math vol. X1V
   p. 355-369 (1961).

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER<br>TR-77-18 | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle)<br>A GENERALIZATION OF THE FRANK-WOLFE THEOREM | | 5. TYPE OF REPORT & PERIOD COVERED<br>Technical Report |
| | | 6. PERFORMING ORG. REPORT NUMBER<br>77-18 |
| 7. AUTHOR(s)<br>Andre F. Perold | | 8. CONTRACT OR GRANT NUMBER(s)<br>NR-047-064 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Department of Operations Research<br>Stanford University<br>Stanford, CA 94305 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Operations Research Program<br>Office of Naval Research<br>Arlington, VA 22217 | | 12. REPORT DATE<br>July 1977 |
| | | 13. NUMBER OF PAGES<br>19 |
| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) | | 15. SECURITY CLASS. (of this report)<br>UNCLASSIFIED |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale;
its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

N00014-75-C-0267
NSF-MCS-76-81259

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Nonlinear Programming          Optimization
Quadratic Programming          Polyhedral Sets
Norm-coercive functions        Concave functions

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Frank-Wolfe Theorem states that a quadratic function, bounded below on a nonempty polyhedral convex set, attains its infimum there. This paper gives sufficient conditions under which a function either attains its infimum on a nonempty polyhedral convex set or is unbounded below on some halfline of that set. Quadratic functions are shown to satisfy these sufficient conditions.

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73
S/N 0102-014-6601